

6/4



OIKE

RAW SEQUENCE LISTING

DATE: 06/04/2002

PATENT APPLICATION: US/10/072,130

TIME: 11:01:31

Input Set : N:\Crf3\RULE60\10072130.raw

Output Set: N:\CRF3\06042002\J072130.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Bandman, Olga
 6 Goli, Surya K.
 7 Lal, Preeti
 8 Corley, Neil C.
 9 Zhang, Hong

11 (ii) TITLE OF INVENTION: NEW PROTEIN PHOSPHATASE

13 (iii) NUMBER OF SEQUENCES: 4

15 (iv) CORRESPONDENCE ADDRESS:

16 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
 17 (B) STREET: 3174 Porter Drive
 18 (C) CITY: Palo Alto
 19 (D) STATE: CA
 20 (E) COUNTRY: USA
 21 (F) ZIP: 94304

23 (v) COMPUTER READABLE FORM:

24 (A) MEDIUM TYPE: Diskette
 25 (B) COMPUTER: IBM Compatible
 26 (C) OPERATING SYSTEM: DOS
 27 (D) SOFTWARE: FastSEQ for Windows Version 2.0

29 (vi) CURRENT APPLICATION DATA:

C--> 30 (A) APPLICATION NUMBER: US/10/072,130
 C--> 31 (B) FILING DATE: 05-Feb-2002

33 (vii) PRIOR APPLICATION DATA:

34 (A) APPLICATION NUMBER: US/08/873,093
 35 (B) FILING DATE:

37 (viii) ATTORNEY/AGENT INFORMATION:

38 (A) NAME: Billings, Lucy J.
 39 (B) REGISTRATION NUMBER: 36,749
 40 (C) REFERENCE/DOCKET NUMBER: PF-0319 US

42 (ix) TELECOMMUNICATION INFORMATION:

43 (A) TELEPHONE: 415-855-0555
 44 (B) TELEFAX: 415-845-4166

46 (2) INFORMATION FOR SEQ ID NO: 1:

48 (i) SEQUENCE CHARACTERISTICS:

49 (A) LENGTH: 478 amino acids
 50 (B) TYPE: amino acid
 51 (C) STRANDEDNESS: single
 52 (D) TOPOLOGY: linear

54 (vii) IMMEDIATE SOURCE:

55 (A) LIBRARY: THPIPLB01
 56 (B) CLONE: 13177

ENTERED

RAW SEQUENCE LISTING

DATE: 06/04/2002

PATENT APPLICATION: US/10/072,130

TIME: 11:01:31

Input Set : N:\Crf3\RULE60\10072130.raw

Output Set: N:\CRF3\06042002\J072130.raw

```

60      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
62 Met Gly Ala Phe Leu Asp Lys Pro Lys Thr Glu Lys His Asn Ala His
63   1      5      10      15
64 Gly Ala Gly Asn Gly Leu Arg Tyr Gly Leu Ser Ser Met Gln Gly Trp
65      20      25      30
66 Arg Val Glu Met Glu Asp Ala His Thr Ala Val Val Gly Ile Pro His
67      35      40      45
68 Gly Leu Glu Asp Trp Ser Phe Phe Ala Val Tyr Asp Gly His Ala Gly
69      50      55      60
70 Ser Arg Val Ala Asn Tyr Cys Ser Thr His Leu Leu Glu His Ile Thr
71      65      70      75      80
72 Thr Asn Glu Asp Phe Arg Ala Ala Gly Lys Ser Gly Ser Ala Leu Glu
73      85      90      95
74 Leu Ser Val Glu Asn Val Lys Asn Gly Ile Arg Thr Gly Phe Leu Lys
75      100     105     110
76 Ile Asp Glu Tyr Met Arg Asn Phe Ser Asp Leu Arg Asn Gly Met Asp
77      115     120     125
78 Arg Ser Gly Ser Thr Ala Val Gly Val Met Ile Ser Pro Lys His Ile
79      130     135     140
80 Tyr Phe Ile Asn Cys Gly Asp Ser Arg Ala Val Leu Tyr Arg Asn Gly
81      145     150     155     160
82 Gln Val Cys Phe Ser Thr Gln Asp His Lys Pro Cys Asn Pro Arg Glu
83      165     170     175
84 Lys Glu Arg Ile Gln Asn Ala Gly Gly Ser Val Met Ile Gln Arg Val
85      180     185     190
86 Asn Gly Ser Leu Ala Val Ser Arg Ala Leu Gly Asp Tyr Asp Tyr Lys
87      195     200     205
88 Cys Val Asp Gly Lys Gly Pro Thr Glu Gln Leu Val Ser Pro Glu Pro
89      210     215     220
90 Glu Val Tyr Glu Ile Leu Arg Ala Glu Glu Asp Glu Phe Ile Ile Leu
91      225     230     235     240
92 Ala Cys Asp Gly Ile Trp Asp Val Met Ser Asn Glu Glu Leu Cys Glu
93      245     250     255
94 Tyr Val Lys Ser Arg Leu Glu Val Ser Asp Asp Leu Glu Asn Val Cys
95      260     265     270
96 Asn Trp Val Val Asp Thr Cys Leu His Lys Gly Ser Arg Asp Asn Met
97      275     280     285
98 Ser Ile Val Leu Val Cys Phe Ser Asn Ala Pro Lys Val Ser Asp Glu
99      290     295     300
100 Ala Val Lys Lys Asp Ser Glu Leu Asp Lys His Leu Glu Ser Arg Val
101      305     310     315     320
102 Glu Glu Ile Met Glu Lys Ser Gly Glu Glu Gly Met Pro Asp Leu Ala
103      325     330     335
104 His Val Met Arg Ile Leu Ser Ala Glu Asn Ile Pro Asn Leu Pro Pro
105      340     345     350
106 Gly Gly Gly Leu Ala Gly Lys Arg Asn Val Ile Glu Ala Val Tyr Ser
107      355     360     365
108 Arg Leu Asn Pro His Arg Glu Ser Asp Gly Ala Ser Asp Glu Ala Glu
109      370     375     380

```

RAW SEQUENCE LISTING

DATE: 06/04/2002

PATENT APPLICATION: US/10/072,130

TIME: 11:01:31

Input Set : N:\Crf3\RULE60\10072130.raw

Output Set: N:\CRF3\06042002\J072130.raw

```

110 Glu Ser Gly Ser Gln Gly Lys Leu Val Glu Ala Leu Arg Gln Met Arg
111 385                      390                      395                      400
112 Ile Asn His Arg Gly Asn Tyr Arg Gln Leu Leu Glu Glu Met Leu Thr
113                      405                      410                      415
114 Ser Tyr Arg Leu Ala Lys Val Glu Gly Glu Glu Ser Pro Ala Glu Pro
115                      420                      425                      430
116 Ala Ala Thr Ala Thr Ser Ser Asn Ser Asp Ala Gly Asn Pro Thr Met
117                      435                      440                      445
118 Gln Glu Ser His Thr Glu Ser Glu Ser Gly Leu Ala Glu Leu Asp Ser
119                      450                      455                      460
120 Ser Asn Glu Asp Ala Gly Thr Lys Met Ser Gly Glu Lys Ile
121 465                      470                      475
123 (2) INFORMATION FOR SEQ ID NO: 2:
125 (i) SEQUENCE CHARACTERISTICS:
126 (A) LENGTH: 2268 base pairs
127 (B) TYPE: nucleic acid
128 (C) STRANDEDNESS: single
129 (D) TOPOLOGY: linear
131 (vii) IMMEDIATE SOURCE:
132 (A) LIBRARY: 13177
133 (B) CLONE: THPIPLB01
135 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
137 ATATTGTACC TATCAGGCGT CAGCTCTCAA TCTAGATCCC TCCCTGGCCT CGGACTTATT 60
138 GCAAAACATG GGTGCTTTT TGGATAAAC CAAACTGAA AAACATAATG CTCATGGTGC 120
139 TGGGAATGGT TTACGTTATG GCCTGAGCAG CATGCAAGGA TGGAGAGTGG AAATGGAAGA 180
140 TGCACACACA GCTGTTGTAG GTATTCCTCA CGGCTTGGA GACTGGTCAT TTTTTCAGT 240
141 TTATGATGGT CATGCTGGAT CCCGAGTGGC AAATTACTGC TCAACACATT TATTAGAACA 300
142 CATCACTACT AACGAAGACT TTAGGGCAGC TGGAAAATCA GGATCTGCTC TTGAGCTTTC 360
143 AGTGGAAAAAT GTTAAGAATG GTATCAGAAC TGGATTTTTG AAAATTGATG AATACATGCG 420
144 TAACTTTTCA GACCTCAGAA ACGGGATGGA CAGGAGTGGT TCAACTGCAG TGGGAGTTAT 480
145 GATTTACCTT AAGCATATCT ACTTTATCAA CTGTGGTGAT TCACGTGCTG TTCTGTATAG 540
146 GAATGGACAA GTCTGCTTTT CTACCCAGGA TCACAAACCT TGCAATCCAA GGGAAAAGGA 600
147 GCGAATCCAA AATGCAGGAG GCAGCGTGAT GATACAACGT GTTAATGGTT CATTAGCAGT 660
148 ATCTCGTGCT CTGGGGGACT ATGATTACAA GTGTGTTGAT GGCAAGGGCC CAACAGAACA 720
149 ACTTGTTTCT CCAGAGCCTG AGGTTTATGA AATTTTAAGA GCAGAAGAGG ATGAATTTAT 780
150 CATCTTGGCT TGTGATGGGA TCTGGGATGT TATGAGTAAT GAGGAGCTCT GTGAATATGT 840
151 TAAATCTAGG CTTGAGGTAT CTGATGACCT GGAAAATGTG TGCAATTGGG TAGTGGACAC 900
152 TTGTTTACAC AAGGGAAGTC GAGATAACAT GAGTATTGTA CTAGTTTGCT TTCAAATGC 960
153 TCCCAAGGTC TCAGATGAAG CGGTGAAAAA AGATTCAGAG TTGGATAAGC ACTTGAATC 1020
154 ACGGGTTGAA GAGATTATGG AGAAGTCTGG CGAGGAAGGA ATGCCTGATC TTGCCCATGT 1080
155 CATGCGCATC TTGTCTGCAG AAAATATCCC AAATTGCTT CCTGGGGGAG GTCTGCTGG 1140
156 CAAGCGTAAT GTTATTGAAG CTGTTTATAG TAGACTGAAT CCACATAGAG AAAGTGATGG 1200
157 GCCTCCGAT GAAGCAGAGG AAAGTGGATC ACAGGGAAAA TTGGTGGAAG CTCTCAGGCA 1260
158 AATGAGAATT AATCATAGGG GAAACTACCG ACAACTTCTG GAGGAGATGC TGAAGTATTA 1320
159 CAGGCTAGCT AAAGTAGAGG GAGAAGAAAAG CCCTGCTGAA CCAGCTGCCA CAGCTACTTC 1380
160 TTCGAACAGT GATGCTGGAA ACCCAGTGAC AATGCAGGAA AGCCATACTG AATCAGAAAAG 1440
161 TGGTCTTGCT GAATTAGACA GCTCTAATGA AGATGCAGGG ACAAAGATGA GTGGTGAAAA 1500
162 AATATGACTT TCCTTTTTTG TAATATTTTT GTGATCTTTG ATGGTTTTTA ACCTAGGAAG 1560
163 TGTAATGTAT GCATTTATAT AACTGTTTTG TTATTTGAAT CTTGGAAAAC TAGTTTTATT 1620

```

RAW SEQUENCE LISTING

DATE: 06/04/2002

PATENT APPLICATION: US/10/072,130

TIME: 11:01:31

Input Set : N:\Crf3\RULE60\10072130.raw

Output Set: N:\CRF3\06042002\J072130.raw

```

164 ATATTCAGAT AGCCTTGTTT TTTAAAAAGG CCTTTGCATA CACCTTTATG AGATAGTGTA 1680
165 AAATTGACTA TTTATAGTAC TATGGATTTA ATGAAATTAT ATGTCATTTT ACATTGTATG 1740
166 CCAGAAATTA GGCTACCAAT TATGAATTAA AGTCAGTAGT TAAATTAATA CTAGATAGAA 1800
167 TTAGAAATTT TGATTAGAGA GATTATGCTA TATTATGGAA AACTTGTTA ATGTAGAATT 1860
168 ATACTGCTTC ATATTATTTT ACCTATTAGT AACTCATAG TTAGCTTTGT AATAAATTTA 1920
169 TGTTTCTTTT AATAATTTTA GTTCTTCAA GAATGGCTGA TGCTGGCCTG TAATTTTCTT 1980
170 TTCAAGGATG ATAATTGTG TGTGTTTGA TTTGTTTATA TTTTACATCT CTGTAGTTT 2040
171 ATTTTATAGAA GTTGTGAGAT ATTGGATGTG TGGCTATTTT TCCTTTCTCT GTATTCTTTA 2100
172 TGAAACATAA CTTTGGAAAA ACCTATGTAT TATTCATACA GCTTTGGTTT GTATATTCTG 2160
173 TATAGCCTAA CTACACACAT CAAAATGTAT GTCAACCAAG TGTTTAGAAT GAAATTATAA 2220
174 GTGTTTAAAGT CCAAATAAAG CATGTGATGT GGAATAATCA AAAAAAAA 2268

```

176 (2) INFORMATION FOR SEQ ID NO: 3:

178 (i) SEQUENCE CHARACTERISTICS:

179 (A) LENGTH: 390 amino acids

180 (B) TYPE: amino acid

181 (C) STRANDEDNESS: single

182 (D) TOPOLOGY: linear

184 (vii) IMMEDIATE SOURCE:

185 (A) LIBRARY: GenBank

186 (B) CLONE: 1247927

188 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

190 Met Gly Ala Phe Leu Asp Lys Pro Lys Thr Glu Lys His Asn Ala His
191 1 5 10 15
192 Gly Ala Gly Asn Gly Leu Arg Tyr Gly Leu Ser Ser Met Gln Gly Trp
193 20 25 30
194 Arg Val Glu Met Glu Asp Ala His Thr Ala Val Val Gly Ile Pro His
195 35 40 45
196 Gly Leu Glu Asp Trp Ser Phe Phe Ala Val Tyr Asp Gly His Ala Gly
197 50 55 60
198 Ser Arg Val Ala Asn Tyr Cys Ser Thr His Leu Leu Glu His Ile Thr
199 65 70 75 80
200 Thr Asn Glu Asp Phe Arg Ala Ala Asp Lys Ser Gly Phe Ala Leu Glu
201 85 90 95
202 Pro Ser Val Glu Asn Val Lys Thr Gly Ile Arg Thr Gly Phe Leu Lys
203 100 105 110
204 Ile Asp Glu Tyr Met Arg Asn Phe Ser Asp Leu Arg Asn Gly Met Asp
205 115 120 125
206 Arg Ser Gly Ser Thr Ala Val Gly Val Met Ile Ser Pro Thr His Ile
207 130 135 140
208 Tyr Phe Ile Asn Cys Gly Asp Ser Arg Ala Val Leu Cys Arg Asn Gly
209 145 150 155 160
210 Gln Val Cys Phe Ser Thr Gln Asp His Lys Pro Cys Asn Pro Met Glu
211 165 170 175
212 Lys Glu Arg Ile Gln Asn Ala Gly Gly Ser Val Met Ile Gln Arg Val
213 180 185 190
214 Asn Gly Ser Leu Ala Val Ser Arg Ala Leu Gly Asp Tyr Asp Tyr Lys
215 195 200 205
216 Cys Val Asp Gly Lys Gly Pro Thr Glu Gln Leu Val Ser Pro Glu Pro
217 210 215 220

```

RAW SEQUENCE LISTING

DATE: 06/04/2002

PATENT APPLICATION: US/10/072,130

TIME: 11:01:31

Input Set : N:\Crf3\RULE60\10072130.raw

Output Set: N:\CRF3\06042002\J072130.raw

```

218  Glu Val Tyr Glu Ile Leu Arg Ala Glu Glu Asp Glu Phe Val Val Leu
219  225                230                235                240
220  Ala Cys Asp Gly Ile Trp Asp Val Met Ser Asn Glu Glu Leu Cys Glu
221                245                250                255
222  Phe Val Asn Ser Arg Leu Glu Val Ser Asp Asp Leu Glu Asn Val Cys
223                260                265                270
224  Asn Trp Val Val Asp Thr Cys Leu His Lys Gly Ser Arg Asp Asn Met
225                275                280                285
226  Ser Ile Val Leu Val Cys Phe Ala Asn Ala Pro Lys Val Ser Asp Glu
227                290                295                300
228  Ala Val Lys Arg Asp Leu Glu Leu Asp Lys His Leu Glu Ser Arg Val
229  305                310                315                320
230  Glu Glu Ile Met Gln Lys Ser Gly Glu Glu Gly Met Pro Asp Leu Ala
231                325                330                335
232  His Val Met Arg Ile Leu Ser Ala Glu Asn Ile Pro Asn Leu Pro Pro
233                340                345                350
234  Gly Gly Gly Leu Ala Gly Lys Arg Asn Val Ile Glu Ala Val Tyr Ser
235                355                360                365
236  Arg Leu Asn Pro Asn Lys Asp Asn Asp Gly Gly Ala Gly Asp Leu Glu
237                370                375                380
238  Asp Ser Leu Val Ala Leu
239  385                390

```

243 (2) INFORMATION FOR SEQ ID NO: 4:

245 (i) SEQUENCE CHARACTERISTICS:

246 (A) LENGTH: 478 amino acids

247 (B) TYPE: amino acid

248 (C) STRANDEDNESS: single

249 (D) TOPOLOGY: linear

251 (vii) IMMEDIATE SOURCE:

252 (A) LIBRARY: GenBank

253 (B) CLONE: 1452526

255 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

257  Met Gly Ala Phe Leu Asp Lys Pro Lys Thr Glu Lys His Asn Ala His
258  1          5          10          15
259  Gly Ala Gly Asn Gly Leu Arg Tyr Gly Leu Ser Ser Met Gln Gly Trp
260                20                25                30
261  Arg Val Glu Met Glu Asp Ala His Thr Ala Val Val Gly Ile Pro His
262                35                40                45
263  Gly Leu Glu Asp Trp Ser Phe Phe Ala Val Tyr Asp Gly His Ala Gly
264                50                55                60
265  Ser Arg Val Ala Asn Tyr Cys Ser Thr His Leu Leu Glu His Ile Thr
266  65                70                75                80
267  Thr Asn Glu Asp Phe Arg Ala Ala Gly Lys Ser Gly Ser Ala Leu Glu
268                85                90                95
269  Leu Ser Val Glu Asn Val Lys Asn Gly Ile Arg Thr Gly Phe Leu Lys
270                100               105               110
271  Ile Asp Glu Tyr Met Arg Asn Phe Ser Asp Leu Arg Asn Gly Met Asp
272                115               120               125
273  Arg Ser Gly Ser Thr Ala Val Gly Val Met Ile Ser Pro Lys His Ile

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/072,130

DATE: 06/04/2002

TIME: 11:01:32

Input Set : N:\Crf3\RULE60\10072130.raw

Output Set: N:\CRF3\06042002\J072130.raw

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]